CS 294-165 Sketching Algorithms — Fall 2020

Problem Set 2

Due: 11:59pm, Monday November 9th

See homework policy at `http://www.sketchingbigdata.org/fall20/syllabus/`
Each problem is worth 15 points (except Problem 4).

Problem 1: Subspace embeddings from JL. Recall Π is an ε-subspace embedding for E being the column space of $U \in \mathbb{R}^{n\times d}$ (which has orthonormal columns) if $\|M\| \leq \varepsilon$ for $M = (\Pi U)^\top(\Pi U) - I$. Recall also $\|M\|$ for a real symmetric matrix equals $\sup_{\|x\|_2=1} |x^\top M x|$. For S a (possibly infinite) set of vectors, we say Π is an (ε, S)-isometry if

$$\forall z \in S,\ (1-\varepsilon)\|z\|_2^2 \leq \|\Pi z\|_2^2 \leq (1+\varepsilon)\|z\|_2^2.$$

(a) (10 points) Show that for any E as above there exists a set S_E of size $O(1)^d$ (independent of ε) such that if Π is an (ε, S_E)-isometry, then it is an $O(\varepsilon)$-subspace embedding for E. Conclude that OSE's exist with $m = O((d+\log(1/\delta))/\varepsilon^2)$ rows. **Hint:** Consider letting S_E be a γ-net of E for some constant γ; what can you say about the vector x^* which achieves the supremum $|x^\top M x|$?

(b) (5 points) Show that (ε, k) RIP matrices exist (see Definition 5.3.5 of the course notes) with $m = O(\varepsilon^{-2} k \log(2n/k))$ rows. **Hint:** use (a), and you may use without proof the fact that $\log \binom{n}{k} = \Theta(k \log(2n/k))$.

Problem 2: Estimating leverage scores. We saw two ways of spectrally approximating $A^T A$: one was to use an oblivious subspace embedding Π and output $(\Pi A)^T(\Pi A)$, and the other approach was to sample rows of A using leverage scores. The naive way to calculate leverage scores is to compute the SVD of A, $A = U\Sigma V^T$, then the ith leverage score ℓ_i is the squared norm of the ith row of U. Computing the SVD of A is usually what we want to avoid though, in order to speed up algorithms! In this problem, we will see how to quickly obtain approximate leverage scores (note for the applications discussed so far in class, it sufficed to know them just up to a constant factor).

In the problems below, recall that for a matrix A with SVD $U\Sigma V^T$, we define $f(A)$ to be $Uf(\Sigma)V^T$, where if Σ is the diagonal matrix with diagonal entries $\sigma_1, \ldots, \sigma_r$, then $f(\Sigma)$ has diagonal entries $f(\sigma_1), \ldots, f(\sigma_r)$ (we will below mention matrices $(A^T A)^{1/2}$, for example, so $f(x) = \sqrt{x}$). Assume below that A has full column rank.

(a) (4 points) Show that Π being an ε-subspace embedding for the column space of A is equivalent to the condition

$$\|(A^T A)^{-1/2}(\Pi A)^T(\Pi A)(A^T A)^{-1/2} - I\| < \varepsilon,$$

i.e. all eigenvalues of $(A^T A)^{-1/2}(\Pi A)^T(\Pi A)(A^T A)^{-1/2}$ are in $[1-\varepsilon, 1+\varepsilon]$.

(b) (4 points) Use (a) to conclude that if $x^T(\Pi A)^T(\Pi A)x = (1 \pm \varepsilon)x^T A^T Ax$ for all x, then $x^T((\Pi A)^T(\Pi A))^{-1}x = (1 \pm O(\varepsilon))x^T(A^T A)^{-1}x$ for all x.

(c) (7 points) Devise an algorithm to compute all leverage scores $\ell_1, \ldots, \ell_n$ of A up to a constant multiplicative factor (i.e. you should output a sequence $\tilde{\ell}_1, \ldots, \tilde{\ell}_n$ such that $(1/2)\ell_i \leq \tilde{\ell}_i \leq 2\ell_i$ for all i). Your algorithm should take time $O((\text{nnz}(A) + \text{poly}(d)) \log n)$ to receive full credit, where $\text{nnz}(A)$ is the number of nonzero entries of A. We assume here $\text{nnz}(A) \geq n$ (otherwise A has at least one row which is entirely zero). **Hint:** Show $\ell_i = \|A(A^T A)^{-1}a_i\|_2^2$ where a_i is the ith row of A. Replace $A^T A$ with $(\Pi A)^T(\Pi A)$ for an oblivious subspace embedding Π, then use the JL lemma to quickly obtain all the desired squared ℓ_2 norms over each i.

Problem 3: (1 point) How much time did you spend on this problem set? If you can remember the breakdown, please report this per problem. (sum of time spent solving problem and typing up your solution)